



# UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE  
United States Patent and Trademark Office  
Address: COMMISSIONER FOR PATENTS  
P.O. Box 1450  
Alexandria, Virginia 22313-1450  
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/005,875	12/07/2001	Stephen John Hinde	1509-253	7959

22429 7590 11/26/2004

LOWE HAUPTMAN GILMAN AND BERNER, LLP  
1700 DIAGONAL ROAD  
SUITE 300 /310  
ALEXANDRIA, VA 22314

EXAMINER

ALBERTALLI, BRIAN LOUIS

ART UNIT PAPER NUMBER

2655

DATE MAILED: 11/26/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

## Office Action Summary

Application No.

10/005,875

Applicant(s)

HINDE ET AL.

Examiner

Brian L Albertalli

Art Unit

2655

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

### Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

### Status

- 1) ☐ Responsive to communication(s) filed on \_\_\_\_.
- 2a) ☐ This action is **FINAL**. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

### Disposition of Claims

- 4) ☒ Claim(s) 1-29 is/are pending in the application.
- 4a) Of the above claim(s) \_\_\_\_ is/are withdrawn from consideration.
- 5) ☐ Claim(s) \_\_\_\_ is/are allowed.
- 6) ☒ Claim(s) 1-29 is/are rejected.
- 7) ☐ Claim(s) \_\_\_\_ is/are objected to.
- 8) ☐ Claim(s) \_\_\_\_ are subject to restriction and/or election requirement.

### Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on \_\_\_\_ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.  
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).  
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

### Priority under 35 U.S.C. § 119

- 12) ☒ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☒ All b) ☐ Some \* c) ☐ None of:
1. ☒ Certified copies of the priority documents have been received.
  2. ☐ Certified copies of the priority documents have been received in Application No. \_\_\_\_.
  3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

\* See the attached detailed Office action for a list of the certified copies not received.

### Attachment(s)

- |  |   |
|--|---|
| 1) <input type="checkbox"/> Notice of References Cited (PTO-892)                                   | 4) <input type="checkbox"/> Interview Summary (PTO-413)                     |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948)               | Paper No(s)/Mail Date. ____   |
| 3) <input checked="" type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| Paper No(s)/Mail Date <u>12/7/01, 3/4/02</u>   | 6) <input type="checkbox"/> Other: ____                                     |

## DETAILED ACTION

### *Claim Objections*

1. Claim 3 is objected to because of the following informalities: in line 1, "are" should be --is--. Appropriate correction is required.

### *Claim Rejections - 35 USC § 102*

2. The following is a quotation of the appropriate paragraphs of 35 U.S.C. 102 that form the basis for the rejections under this section made in this Office action:

A person shall be entitled to a patent unless –

(b) the invention was patented or described in a printed publication in this or a foreign country or in public use or on sale in this country, more than one year prior to the date of application for patent in the United States.

3. Claims 1-7, 14-16, and 18-22 are rejected under 35 U.S.C. 102(b) as being anticipated by Rtischev et al. (U.S. Patent 5,634,086).

In regard to claim 1 and 18, Rtischev et al. discloses a method (Fig. 4A1 and Fig. 4A2) of interacting with a human user through a sound service system, wherein the service system participates with the human user both in normal voice dialog exchanges (steps G-AB), and in a multi-turn sound exchange (steps B-F) the form and content of which are pre-specified and already public (Fig. 1, remote user 12' employs well-known text as prompts to the instructional apparatus 10, column 4, lines 48-50), this sound exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance with the appropriate pre-specified content.

In recognition mode, the instructional apparatus 10 presents a pre-selected script to a user whose pronunciation is to be evaluated (column 5, lines 47-51). As the remote

Art Unit: 2655

user 12' reads the pre-selected, well-known text, a tracking process is performed to determine whether the user has read the sentence satisfactorily (column 6, lines 14-21). If the words read were appropriate (good), the instructional apparatus 10 returns the pre-specified response, "okay" (column 6, lines 21-24), and continues the multi-turn sound exchange.

In regard to claim 2 and 20, Rtischev et al. discloses multi-turn sound exchange serves no function in respect of restricting access to, or controlling the course of, the normal dialog exchanges.

As long as the user continues to speak an appropriate (good) response, the multi-turn dialog continues (steps B-F, column 6, lines 14-24). The multi-turn exchange (steps B-F) therefore serves no function in respect of controlling the course of the normal dialog exchanges (steps G-AB). Furthermore, the multi-turn exchange (steps B-F) serves no function in respect of restricting access to the normal dialog exchanges (steps G-AB), since any inaccurate (not good) response will serve to access the normal dialog exchanges (steps G-AB, column 6, lines 25-28).

In regard to claim 3 and 21, Rtischev et al. discloses the multi-turn sound exchanges are of a promotional nature (the prompts are read from a newspaper advertisement, column 4, line 50).

Art Unit: 2655

In regard to claim 4, Rtischev et al. discloses the multi-turn sound exchange is initiated by the service system (steps Z-AB, if good words are not detected, the system loses patience and initiates the multi-turn sound exchange at step B).

In regard to claim 5, Rtischev et al. discloses the multi-turn sound exchange is initiated by the human user (initially, at step B the multi-turn sound exchange (steps B-F) is initiated by the user, column 6, lines 12-16).

In regard to claim 6, Rtischev et al. discloses the multi-turn sound exchange is initiated at any time during the course of the normal dialog exchanges (as soon as the user speaks a good sentence and continues to a next sentence, the multi-turn sound exchange (steps B-F) is initiated at steps M and V, column 6, lines 39-44 and lines 58-62).

In regard to claim 7, Rtischev et al. discloses the service system uses the same dialog manager (application subsystem) for the normal voice dialogs and the multi-turn sound exchanges with each being effected according to a corresponding script run by the dialog manager as required (Fig. 2, application subsystem 48 controls the interaction of the user 12 and the lesson material for all steps in Fig. 4A and Fig. 4B, column 6, lines 8-11).

Art Unit: 2655

In regard to claim 14, Rtischev et al. discloses the multi-turn sound exchange includes non-word sounds (step E checks for pausing by the user, column 6, lines 18-21, <sup>background noise</sup> pausing is a non-word sound, column 8, lines 40-44).

In regard to claim 15, Rtischev et al. discloses the multi-turn sound exchange is of a looping nature (see Fig. 4A, steps B-F are looping) and terminates in response to at least one of:

explicit user request (at step E, a speaker can request to terminate the multi-turn sound exchange (steps B-F) by speaking unrecognizable words or not pausing after reading good words, column 6, lines 18-21 and lines 25-27);

execution of a preset number of cycles (at step C if the user has completed reading the previously known script, the multi-turn sound exchange (steps B-F) terminates at step D, column 6, lines 16-18).

In regard to claim 16, Rtischev et al. discloses the user's input during at least one turn of the multi-turn sound exchange, is used to determine which of two or more branches in the service system's part of the multi-turn sound exchange is taken by the service system (one at step C, to determine whether the last sentence was input by the user, column 6, lines 16-18; and one at step E, if the user has given a good input, the multi-turn sound exchange will branch to the next sentence, column 6, lines 21-24, if not, the multi-turn sound exchange will branch to the normal voice dialog at step G, column 6, lines 25-28).

In regard to claim 19, Rtischev et al. discloses a sound service system (Fig. 1, 10) comprising a sound input channel for receiving (telephone 14, telephone network 24, MUX 26, preamp 28, low pass filter 30 and A/D 32) and interpreting (speech recognition system controlled by DSP 34 and workstation 36) sound input signals, a sound output channel for generating sound output signals (D/A 38, telephone network 24, and telephone 14), and a dialog manager (application subsystem 48) connected to an output of the sound input channel and an input of the sound output channel, the dialog manager being operative to manage the participation of the service system in exchanges with a user (application subsystem 48 is run on workstation 36 and controls the interaction of the user 12 and lesson material, column 5, lines 4-8 and column 6, lines 8-11) and comprising:

means for managing participation of the service system in normal voice dialog exchanges with the user (application subsystem 48 controls the interaction of the user 12, including the normal voice dialog, steps G-AB, column 6, lines 8-11), and

means for managing participation of the service system in a multi-turn sound exchange with the user, the form and content of this exchange being pre-specified and already public, and the exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance with the appropriate pre-specified content.

Application subsystem 48 additionally controls the multi-turn sound exchange (steps B-F, column 6, lines 8-11). In recognition mode, the instructional apparatus 10

Art Unit: 2655

presents a pre-selected script to a user whose pronunciation is to be evaluated (column 5, lines 47-51). As the remote user 12' reads the pre-selected, well-known text, a tracking process is performed to determine whether the user has read the sentence satisfactorily (column 6, lines 14-21). If the words read were appropriate (good), the instructional apparatus 10 returns the pre-specified response, "okay" (column 6, lines 21-24), and continues the multi-turn sound exchange.

In regard to claim 22, Rtischev et al. discloses the dialog manager includes initiation means for initiating a multi-turn sound exchange under the control of the corresponding said means for managing, the initiation means being operative to initiate a multi-turn sound exchange in response to an input by the human user made at any time during the course of a said normal voice dialog exchange.

In steps Z-AB, if good words are not detected, the system loses patience and initiates the multi-turn sound exchange at step B. Steps Z-AB are controlled by the dialog manager (application subsystem 48), which provides the means for initiating the multi-turn sound exchange at step AB.

### ***Claim Rejections - 35 USC § 103***

The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.



4. Claims 8-12 and 23-29 are rejected under 35 U.S.C. 103(a) as being unpatentable over Rtischev et al., in view of Brown et al. (U.S. Patent 6,377,922).

In regard to claim 8, Rtischev et al. discloses that the dialog control for the multi-turn sound exchanges (check for good words tracked at step E) is separate from the dialog control for the normal voice dialogs (check for good words tracked at step K and T). Furthermore, Rtischev et al. discloses that the reject indicator generated during tracking is adjustable automatically by the system (column 7, lines 45-49).

Rtischev et al. does not disclose that a separate manager is used for each of the multi-turn sound exchanges and the normal voice dialogs.

Brown et al. discloses a system (Fig. 1, 100) for use over a telephone network that includes multiple managers (speech recognizers 105, 106, and 107) and a switch (104) for passing control between the managers (column 2, lines 64-67). Each of the speech recognizers 105-107 has different capabilities (column 3, lines 23-26).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Rtischev et al. to include two managers, one for the normal voice dialogs and one for the multi-turn sound exchanges, each manager when in control effecting the control according to a corresponding script, so that a separate reject indicator could be generated during tracking (recognition) for both the normal dialogs and the multi-turn sound exchanges. This would serve to increase the accuracy of spoken utterances from a user, as disclosed by Brown et al. (column 6, lines 29-31).

In regard to claim 9, Rtischev et al. discloses the step of a user inputting a sound corresponding to the start of a particular multi-turn sound exchange whilst in the voice dialog (step K, the user reads good words with the appropriate pause, column 6, lines 36-39) and running the script corresponding to said particular multi-turn sound exchange (returning to the multi-turn sound exchange at step B, column 6, lines 39-44).

Furthermore, the combination of Rtischev et al. and Brown et al. as applied to claim 8, above, would necessarily turn control of the dialog from the voice dialog manager to the multi-turn sound exchange manager when the user input the appropriate sound.

In regard to claim 10, Rtischev et al. discloses the service system is adapted to recognize and distinguish between sounds corresponding to multiple different multi-turn sound exchanges (different sound exchanges that the user can read include published or printed text, well-known text or memorized text, column 4, lines 48-50).

In regard to claim 11, Rtischev et al. discloses the step of a user inputting a sound, whilst in a multi-turn dialog indicative that the user wants to exit the current multi-turn sound exchange (step E, a user reads words that are not recognizable or does not pause, column 6, lines 18-21), the service system recognizing the sound and running the appropriate dialog script (if the words are not recognizable or the user does not pause, step G executes, the beginning of the normal voice script, column 6, lines 25-30).

In regard to claim 12, the combination of Rtischev et al. and Brown et al., as applied to claim 8 above, discloses in Brown et al. that the scripts (prompts) for the voice dialog manager and the multi-turn dialog manager are independently loaded (prompts in the database are each associated with one of the recognizers 105-107 and are only loaded for the corresponding recognizer, column 3, lines 60-65).

Furthermore, the combination of Rtischev et al. and Brown et al. as applied to claim 8, above, would necessarily turn control of the dialog from the multi-turn sound exchange manager to the voice dialog manager when the user input the appropriate sound.

In regard to claim 23, Rtischev et al. discloses:

a sound input channel for receiving (telephone 14, telephone network 24, MUX 26, preamp 28, low pass filter 30 and A/D 32) and interpreting (speech recognition system controlled by DSP 34 and workstation 36) sound input signals;

a sound output channel for generating sound output signals (D/A 38, telephone network 24, and telephone 14); and

a dialog manager (application subsystem 48) connected to an output of the sound input channel and an input of the sound output channel, the dialog manager being operative to manage the participation of the service system in exchanges with a user (application subsystem 48 is run on workstation 36 and controls the interaction of

Art Unit: 2655

the user 12 and lesson material, column 5, lines 4-8 and column 6, lines 8-11) and comprising:

means for managing participation of the service system in normal voice dialog exchanges with the user (application subsystem 48 controls the interaction of the user 12, including the normal voice dialog, steps G-AB, column 6, lines 8-11), and

means for managing participation of the service system in a multi-turn sound exchange with the user, the form and content of this exchange being pre-specified and already public, and the exchange involving one or more cycles in each of which the service and user take turns to provide a noise or utterance with the appropriate pre-specified content.

Rtischev et al. does not disclose a separate voice service manager and multi-turn dialog manager or a changeover controller for switching control between the voice service manager and the multi-turn dialog manager.

Brown et al. discloses a system (Fig. 1, 100) for use over a telephone network that includes multiple managers (speech recognizers 105, 106, and 107) and a switch (104) for passing control between the managers (column 2, lines 64-67). Each of the speech recognizers 105-107 has different capabilities (column 3, lines 23-26).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Rtischev et al. to include two managers, one for the normal voice dialogs and one for the multi-turn sound exchanges, and a changeover controller for switching control between the two, so that a separate reject indicator could be generated during tracking (recognition) for both the normal dialogs and the multi-turn

Art Unit: 2655

sound exchanges. This would serve to increase the accuracy of spoken utterances from a user, as disclosed by Brown et al. (column 6, lines 29-31).

In regard to claim 24, Rtischev et al. discloses multi-turn sound exchange serves no function in respect of restricting access to, or controlling the course of, the normal dialog exchanges.

As long as the user continues to speak an appropriate (good) response, the multi-turn dialog continues (steps B-F, column 6, lines 14-24). The multi-turn exchange (steps B-F) therefore serves no function in respect of controlling the course of the normal dialog exchanges (steps G-AB). Furthermore, the multi-turn exchange (steps B-F) serves no function in respect of restricting access to the normal dialog exchanges (steps G-AB), since any inaccurate (not good) response will serve to access the normal dialog exchanges (steps G-AB, column 6, lines 25-28).

In regard to claim 25, Rtischev et al. discloses the multi-turn sound exchanges are of a promotional nature (the prompts are read from a newspaper advertisement, column 4, line 50).

In regard to claim 26, Rtischev et al. discloses the step of a user inputting a sound corresponding to the start of a particular multi-turn sound exchange whilst in the voice dialog (step K, the user reads good words with the appropriate pause, column 6,

Art Unit: 2655

lines 36-39) and running the script corresponding to said particular multi-turn sound exchange (returning to the multi-turn sound exchange at step B, column 6, lines 39-44).

Furthermore, the combination of Rtischev et al. and Brown et al. as applied to claim 23, above, switch 104 would necessarily turn control of the dialog from the voice dialog manager to the multi-turn sound exchange manager when the user input the appropriate sound.

In regard to claim 27, Rtischev et al. discloses the service system is adapted to recognize and distinguish between sounds corresponding to multiple different multi-turn sound exchanges (different sound exchanges that the user can read include published or printed text, well-known text or memorized text, column 4, lines 48-50).

In regard to claim 28, Rtischev et al. discloses the step of a user inputting a sound, whilst in a multi-turn dialog indicative that the user wants to exit the current multi-turn sound exchange (step E, a user reads words that are not recognizable or does not pause, column 6, lines 18-21), the service system recognizing the sound and running the appropriate dialog script (if the words are not recognizable or the user does not pause, step G executes, the beginning of the normal voice script, column 6, lines 25-30).

Furthermore, the combination of Rtischev et al. and Brown et al. as applied to claim 23, above, switch 104 would necessarily turn control of the dialog from the multi-

turn sound exchange manager to the voice dialog manager when the user input the appropriate sound.

In regard to claim 29, Rtischev et al. discloses the multi-turn sound exchange includes non-word sounds (step E checks for pausing by the user, column 6, lines 18-21, pausing is a non-word sound, column 8, lines 40-44), the system including specific means for recognizing and/or generating said non word sounds (Fig. 2, HMM search 44 searches HMM models 46, the models including pause models, column 5, lines 12-17).

5. Claim 13 is rejected under 35 U.S.C. 103(a) as being unpatentable over Rtischev et al., in view of Brown et al., and further in view of VoiceXML Forum (*Voice eXtensible Markup Language*).

Rtischev et al. discloses the voice service system (10) comprises a server (workstation 36) with one or more multi-turn sound exchanges scripts (column 6, lines 12-14). Rtischev et al. further discloses that the script is loaded before the user begins to read (the script is started at sentence index  $i=1$  and word index  $j=1$  in step A before tracking the user's input at step B, column 6, lines 14-16).

Rtischev et al. does not disclose that the voice service system comprises a voice browser for interpreting scripts provided by voice sites hosted by page servers, where the scripts are loaded upon first contact of the voice site and remain loaded whilst the user browses the voice pages of the site, the currently visited voice page of the site being loaded to the voice dialog manager.

VoiceXML discloses a voice browser (implementation platform) for interpreting scripts provided by voice sites (VoiceXML documents) hosted by page servers (document servers, page 7, section 2.1 paragraphs 1 and 2 and Fig. 1). VoiceXML further discloses that loading (caching) improves the performance in fetching documents and other resources in a voice browser (page 42, section 12.2, lines 1-3).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Rtischev et al. and Brown et al. to be implemented as a voice browser that accessed voice sites hosted by servers, since VoiceXML frees the authors of voice response applications from low-level programming and resource management, as taught by VoiceXML (page 8, section 2.2, lines 1-3). Furthermore, it would have been obvious to one of ordinary skill in the art at the time of invention to load the multi-turn sound exchange scripts upon a user first contacting a voice site and to load the voice page of the site to the voice dialog manager, since loading (caching) a voice site improves the performance of fetching documents in a voice service system, as taught by VoiceXML (page 42, section 12.2, lines 1-3).

### ***Conclusion***

6. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Wasowicz (U.S. Patent 6,755,657) discloses a reading training



Art Unit: 2655

system that has the user read back previously known material. McIlwaine et al. (U.S. Patent 6,628,777) discloses a system for distributing previously known interactive training to call center agents. Rigsby et al. (U.S. Patent 6,556,971) discloses a method to associate any sound, including non-speech sounds, with an icon for later recognition. Adams, Jr. et al. (U.S. Patent 6,017,219) discloses a system for a user to read back a previously known script that includes multiple branches based on the user's input. Machin et al. (U.S. Patent 6,038,544) discloses a system for testing a call center agent that presents a prerecorded script to the agent and allows the user to speak back a pre-selected response. Scott et al. (U.S. Patent 4,468,204) discloses a system that records a series of questions and answers that a user student can then be tested on through speech recognition. Blackmer et al. (U.S. Patent 5,393,236) discloses a system for presenting a prerecorded lesson plan for a user to practice pronunciation.

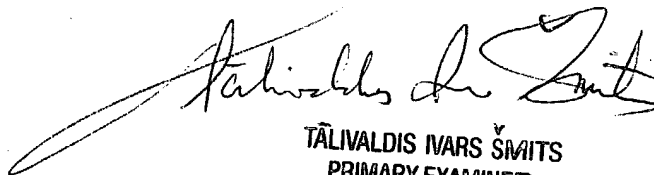
7. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Mon - Fri, 8:00 AM - 5:30 PM, every second Fri off.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Art Unit: 2655

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 11/22/04



TĀIVALDIS IVARS ŠMITS  
PRIMARY EXAMINER